

粗略群集法之屬性選取

許志華

中國文化大學應用數學系

摘要

在粗略集中屬性篩選的技術僅適用於離散型的資料，然而在資訊系統中大多數資料屬性都是連續值。因此，一個結合群聚法和粗略集理論的全域資料離散化和屬性篩選的演算法被提出，稱為粗略 k -中心群集演算法(rough k -means clustering algorithm)。演算法進行資料分析之前必須先決定群集的個數，為了避免此缺點，本論文中，提出一個驗證群集個數是否適當的指標，並舉一關於自我品牌服飾消費者滿意度調查的實際案例說明所提方法的有幸性，實驗結果證明本文中所提的方法是有效的。

關鍵字：資料離散化、屬性篩選、粗略 k -中心群集演算法